Evidence for the Efficacy of Over-the-counter Vitamin A Cosmetic Products in the Improvement of Facial Skin Aging: A Systematic Review

by NATALIA M. K. SPIERINGS, BSC (HONS), MBBS, MRCP (UK), MRCP (DERM), MBA, MSC
Dr. Spierings is with London Medical in London, England and Kings College Hospital London in Dubai, United Arab Emirates.

J Clin Aesthet Dermatol. 2021;14(9):33–40.

ABSTRACT

BACKGROUND: Skincare retailers sell a plethora of retinol-containing products, ranging from serums and moisturisers to masks and eye creams. **OBJECTIVE:** The purpose of this review is to critically appraise the randomized, double-blind, vehicle-controlled trials of the use of over-the-counter retinol products in the treatment of facial skin aging in order to assess evidence regarding their efficacy. **METHODS:** A PubMed search was conducted for relevant clinical trial publications, using the terms "retinoid," "tretinoin," "retinol," "retinal," "retinaldehyde," and "skin." **RESULTS:** Nine randomized, double-blind, vehicle-controlled clinical trials were found. Four of these trials reported no statistically significant differences between the retinol-containing treatment and vehicle. The remaining five trials provide weak evidence for retinol potentially having a mild ameliorating effect on fine facial skin wrinkle lines only. However, these five trials showed major methodological flaws, which were critically analyzed in this review, calling into question the validity of any positive results. **CONCLUSION:** It can be suggested that, in the case of retinols, the "positive" trials should not inform clinical decision-making but rather may serve as tools for advertising. Until at least one high-quality clinical trial of retinol-containing products in the treatment of (photo-)aged skin is published, there is very little, if any, trustworthy evidence available to support the use of over-the-counter cosmetic retinol-containing products to improve the appearance of aged skin.

KEY WORDS: Tretinoin, facial skin, aging, photo-aging, anti-aging, skincare, retinol, wrinkles, vitamin A

The term "retinoid" encompasses a family of chemical compounds that are natural or synthetic derivatives of vitamin A, sharing structural and functional similarities with this vitamin. In the body, retinol is the main circulating form of vitamin A, retinoic acid is its main active metabolite, and the vitamin is stored in the liver in a variety of retinyl ester forms. The structural differences between the compounds are accounted for by the polar end-group, which is hydroxyl in retinols, aldehyde in retinals, carboxylic acid in retinoic acid, and a variety of esters in the retinyl esters, respectively.[1] Retinoic acid, the biologically active form of vitamin A, is converted from retinol and retinyl esters through a number of enzymatic steps, involving oxidation and hydroxylation.[2] A simplified illustration of retinoid metabolism is shown in Figure 1.

More than 2,000 derivatives of vitamin A have been synthesized since 1955, classified into three generations.[1] The first generation includes the naturally occurring, nonaromatic retinoids, which include retinol, retinal, isotretinoin, and tretinoin. The second-generation retinoids are the mono-aromatic compounds, including etretinate and acitretin. Finally, the third-generation retinoids are formed by cyclization of the polyene side chain and include adapalene and tazarotene; these molecules have a more rigid structure than the first- and second-generation retinoids, which makes them bind to a narrower set of receptors.[1]

Tretinoin is a synthetic retinoic acid and is the retinoid most extensively investigated in the treatment of facial skin (photo-)aging.[3] For the sake of clarity, in this paper, the term "tretinoin" will be used to refer to the biologically active retinoid approved by the United States Food and Drug Administration (FDA), while "retinol" will be the term used here for nonprescription vitamin A derivatives, commonly found in cosmetic skincare products, with the ones most often used listed in Table 2. The purpose of the many chemical alterations of the retinol molecule concerns mostly attempts to increase the stability of the compound.

Active retinoic acid is the best-known category of facial skin anti-aging treatment available today. There is high-level scientific evidence, dating back decades, showing the histological effects of tretinoin on collagen synthesis, fibroblast activity, and the inhibition of matrix metalloproteinases. However, there are few high-quality studies providing support for the benefit of these products as anti-aging treatments, despite various retinoic acid

FUNDING: No funding was provided for this article.
DISCLOSURES: The author reports no conflicts of interest relevant to the content of this article.
CORRESPONDENCE: Natalia M. K. Spierings, Bsc (Hons), MBBS, MRCP (UK), MRCP (Derm), MBA, MSc; Email: nspierings@doctors.org.uk

TABLE 1. Definitions of commonly used terms in relation to the use of vitamin A for dermatological purposes

Retinoid	Umbrella term for a family of chemical compounds that are natural or synthetic derivatives of vitamin A. The compounds share structural and functional similarities with vitamin A; all vitamin A derivatives fall under this term. In the author's clinical practice, the term "retinoid" is used specifically to refer to retinoic acid.
Retinoic acid	A biologically active retinoid that acts on specific retinoic acid receptors in human tissue. There are many synthetic versions of retinoic acid, which are used for a variety of dermatological indications, both orally and topically.
Retinol	The main circulating form of vitamin A, which requires conversion to the active retinoic acid to exert a biological effect. In the author's clinical practice and in this paper, all non–biologically active "downstream" vitamin A derivatives are referred to collectively as "retinols." They are often used as ingredients in over-the-counter cosmetic skincare products and are not classified as prescription medicines.
Tretinoin	A synthetic version of biologically active retinoic acid and the most investigated retinoid in the treatment of facial skin (photo-)aging. In many countries, it is available by prescription only, with brand names including, among others, Acretin®, Renova®, Retin-A®, Avita®, Atralin®, Retin-A Micro®, and Avage®.
Adapalene	A third-generation retinoid specifically licensed for the treatment of acne. It selectively binds the retinoic acid beta- and gamma-receptors; brand names include, among others, Differin® and Adaferin®.
Tazarotene	A third-generation retinoid licensed for the treatment of psoriasis, acne, and photo-damage. It selectively binds to the retinoic acid beta- and gamma-receptors; brand names include, among others, Tazorac®, Avage®, and Zorac®.

TABLE 2. Nonprescription vitamin A derivatives commonly used in cosmetic skincare products, which are referred to in this paper collectively as "retinols."

Retinol
Retinaldehyde = retinal
Retinyl esters, e.g., of palmitic acid, oleic acid, stearic acid, or linoleic acid
Retinyl palmitate
Retinyl propionate
Retinyl retinoate
Retinyl N-formyl aspartamate

FIGURE 1. Simplified illustration of the retinoid metabolism. Retinoic acid, a biologically active vitamin A derivative, is formed by the oxidation of retinol via retinaldehyde or directly from retinal. Retinyl esters, the storage form of vitamin A, are converted via hydroxylation to retinol.

formulations being approved by the FDA for the treatment of fine facial wrinkle lines.

In 2000, Renova® (tretinoin 0.02%) received FDA approval for the "mitigation of fine facial wrinkles in patients who use comprehensive skincare and sunlight avoidance programmes." The FDA approved labeling states that "Renova®…does not eliminate wrinkles, repair sun-damaged skin, reverse photoaging, or restore more youthful or younger skin." Indeed, out of the five clinical trials submitted to the FDA, only two showed a statistically significant improvement in the appearance of fine lines, while none demonstrated efficacy for any other signs of facial skin (photo-) aging.[4]

More recently conducted clinical trials have also failed to show benefit on signs of facial skin (photo-)aging other than fine wrinkles. For example, Kang et al.[5] performed a two-year, multicenter, randomized, double-blind, vehicle-controlled trial of the efficacy of tretinoin 0.05% emollient cream versus vehicle in the treatment of photodamaged facial skin. The sample size was based on a power calculation to ensure that estimates of treatment differences would be clinically meaningful. An intention-to-treat analysis of the results was performed, including all randomized subjects, and the Bonferroni–Holms procedure was used to correct for multiple comparisons. The primary efficacy parameters were fine wrinkling, mottled hyperpigmentation, and tactile roughness of the face, rated using a scoring system of 0 to 9, expressed as the percentage of change from baseline.

In the trial and after two years of treatment, tretinoin 0.05% was statistically significantly superior to vehicle in terms of the percentage of subjects who experienced improvement from baseline in fine wrinkling (76% vs. 55%; p=0.002) and mottled hyperpigmentation (74% vs. 65%; p=0.002) but not tactile roughness (76% vs. 65%; p=0.645). It is important to note that, in the placebo group, the percentage of subjects with improvement from baseline for all three primary efficacy parameters was well over 50%, with the caveat that the definition of improvement was not provided.

Retinols are not classified as medicines due to a lack of safety concerns and are commonly found in over-the-counter cosmetic skincare products. They require conversion to active retinoic acid in the body to demonstrate activity[6] and, *in vitro*, have been shown to be 20 times less potent than retinoic acid.[7] Evidence suggests that, once metabolised in the skin, topically applied retinol or retinyl esters are no longer active or are less active than retinol itself.[8] Retinyl palmitate is the most commonly used retinol in cosmetic skincare products but is theoretically the least effective because it requires conversion to active retinoic acid by a two-step oxidative process, which is limited in the skin.[9]

If manufactured under inert atmospheric conditions and stored in aluminium tubes at less than 20°C, retinol can be stable in a formulation for more than six months.[8] However, cosmetic skincare products containing retinol are generally not manufactured or stored in this way, therefore making it likely that the retinol in these products has very little, if any, effect when applied to the skin. The impact of the vehicle on penetration and stability of the retinol in a formulation is also crucial in this regard.[3]

Regulations surrounding the use of retinol in cosmetic skincare products do not require evidence of efficacy because "cosmetics are not medicines."[10] The Cosmetic, Toiletry, and Perfumery Association (CTPA) in the United Kingdom published a guidance document, titled "Confidence in Cosmetic Claims." Its use is intended for industry, explaining that "claims can be substantiated by using either:

formula, experimental studies, consumer perception tests and/or published information or a combination of these." Therefore, companies can make statements about their retinol-containing products like "this intensive treatment minimises visibility of fine lines and dark spots. . . promoting skin renewal and boosting collagen production"[11] without providing evidence of efficacy, presumably basing such claims on the what is known about tretinoin in general.

Online skincare retailers sell a plethora of retinol-containing products, ranging from serums and moisturizers to masks and eye creams. Performing a search for "retinol" on a popular beauty website generates more than 250 products, ranging in price from £4.20 for a 30-mL serum to £210 for a 50-mL cream. The question that many consumers rightly ask themselves is: are these products at all effective? Considering that many of the products are more expensive than the equivalent amount of prescription tretinoin, sometimes, even when adding in the cost of seeing a dermatologist to get the prescription, is it worth purchasing them instead of obtaining a tretinoin prescription?

The purpose of this review is to critically appraise the existing randomized, double-blind, vehicle-controlled trials of the use of over-the-counter cosmetic retinol products in the treatment of facial skin aging in order to assess the level of evidence regarding their efficacy.

METHODS

A PubMed® search was conducted for relevant clinical trial publications, using the terms "retinoid," "tretinoin," "retinol," "retinal," "retinaldehyde," and "skin" but not "acne." The selection of trials was based on the following criterion: clinical trials evaluating the efficacy of retinol in the treatment of facial skin (photo-)aging compared to vehicle or placebo published at any point in time. Exclusion criteria included clinical trials looking at nonfacial skin only; trials without a vehicle or placebo control; active-control trials without a noninferiority design; trials involving retinol combination therapies, such as combined with vitamin C or other "novel" medicinal-type ingredients; and trials reporting only on histological parameters. *Ex-vivo* and *in-vitro* studies were excluded as well, and duplicate publications were identified to ensure that trials were only included once.

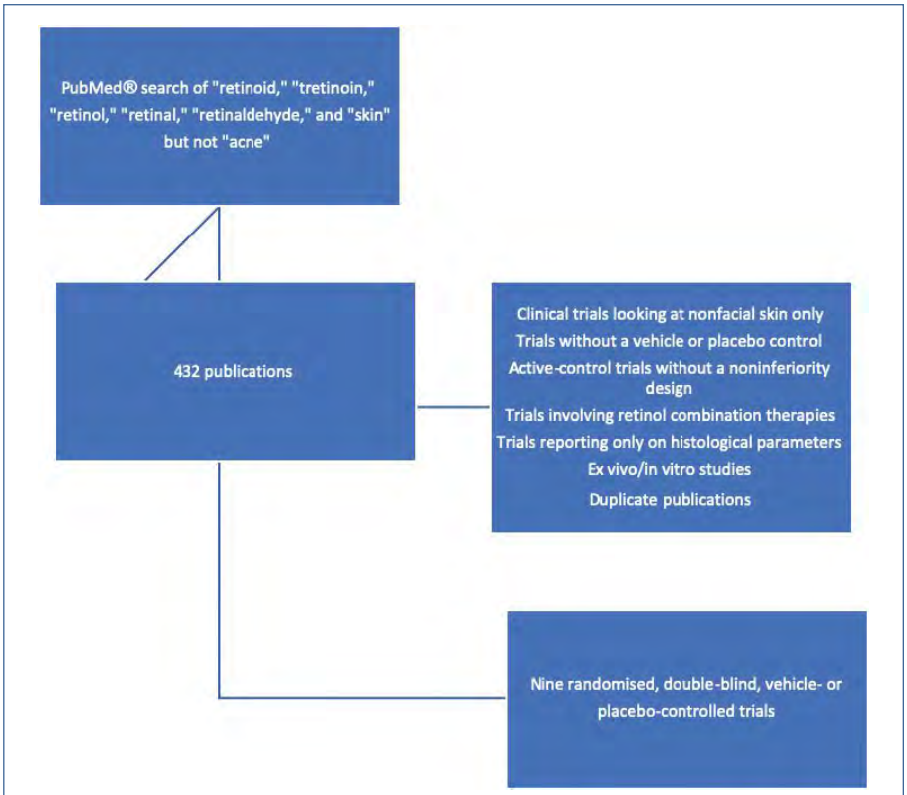

FIGURE 2. A PubMed search was performed for clinical trials evaluating the efficacy of retinol in the treatment of facial skin (photo-)aging relative to vehicle or placebo published at any point in time.

A total of 432 publications were identified and their the titles and abstracts were reviewed for relevants based on the outcome measures most important to patients and clinicians. Subsequently, a total of nine randomized, double-blind, vehicle- or placebo-controlled clinical trials were found based on the above selection criteria (Figure 2), whose quality was assessed using the Jadad criteria.[12]

RESULTS

The nine clinical trials generated by the PubMed search and selected based on the above eligibility and exclusion criteria are summarized in no particular order in Table 3 in terms of intervention, with number of subjects, treatment duration, outcome measures, results, strengths, and weaknesses/limitations reported.

DISCUSSION

Clinical trial methodology is well established in terms of design, execution, analysis, and reporting. Trials should follow the established methodology to allow for accurate and generalizable conclusions, which is important from scientific and clinical perspectives alike. If trials do not follow the established methodology, it becomes difficult to draw meaningful conclusions from them or conclusions that are valid. Some of the key aspects of clinical trial methodology that are relevant here are briefly defined and explained in Table 4.

All nine trials that the search yielded suffer from similar methodological problems. None of the trials had an endpoint designation and, therefore, none had a power calculation as part of the statistical analysis to establish how many subjects would be required for valid conclusions to be drawn. In addition, only one trial (Kim et al.[13]) described the method used to achieve randomization. In two of the trials described as double-blind studies, it was not specifically stated that the investigator was blinded. None of the trials described how, in fact, blinding was achieved and one trial was actually only subject-blinded (Lee et al.[14]),

TABLE 3. Randomized, double-blind*, vehicle-controlled, whole- or split-face design trials of retinols in facial skin aging

STUDY (JADAD SCORE)	INTERVENTION WITH NUMBER OF SUBJECTS	TREATMENT DURATION	OUTCOME MEASURES	RESULTS	STRENGTHS	WEAKNESSES/LIMITATIONS
Creidi et al.[27] (Jadad score: 3)	• Whole-face design • Retinaldehyde 0.05%: 40 subjects • Vehicle: 45 subjects	44 weeks	• Optical profilometry of crow's feet area, assessing for deepness and roughness of fine wrinkles	• No statistically significant differences in retinaldehyde versus vehicle for both outcome measures	• Active comparator with retinoic acid 0.05%: 40 subjects (internal validation) • Objective outcome measure • Demographic data and baseline values provided	• No primary endpoint designation • No power calculation • Completer analysis
Green et al.[28] (Jadad score: 3)	• Whole-face design • 80 subjects enrolled • Retinyl propionate 0.15% • Vehicle	48 weeks	• Clinical assessment, subject self-assessment, photography, and optical profilometry of right crow's feet area	• No statistically significant differences in retinyl propionate versus vehicle for all outcome measures	• Subjective and objective outcome measures • Detailed reporting of outcome results with baseline values	• No primary endpoint designation • No power calculation • No group-size numbers • Completer analysis • No correction for multiple comparisons but alpha set low at 0.005
Tucker-Samaras et al.[21] (Jadad score: 3)	• Split-face design (multiple products tested) • Retinol 0.1%: 37 subjects • Vehicle: 29 subjects	8 weeks	• Clinical assessment with photography of 12 skin parameters, using a scoring system of 0 to 9	• Statistically significant improvements ($p<0.05$) on retinol versus vehicle in 10 of the 12 clinically-assessed skin parameters	• Absolute data as well as absolute and percentage changes from baseline provided	• No primary endpoint designation • No power calculation • Completer analysis • No baseline information • No standard deviations or confidence intervals • No correction for multiple comparisons
Kikuchi et al.[24] (Jadad score: 3)	• Split-face design • Retinol 0.075%/ vehicle: 54 subjects	26 weeks	• Clinical assessment with photography of fine wrinkles, deep wrinkles, and pigmentation • Rating scale: 0, no change; 1, slightly better; 2, better; and 3, far better	• Statistically significant improvements ($p<0.05$) in retinol versus vehicle in fine wrinkles as of Week 4 and in deep wrinkles as of Week 26	• Results reported in detail, including efficacy and tolerability (i.e., erythema, swelling, scaling)	• No primary endpoint • No power calculation • No baseline information • Completer analysis • Power artificially increased by analysing observations rather than subjects • No correction for multiple comparisons
Bellemère et al.[20] (Jadad score: 3)	• Whole-face design • Retinol 0.1%: 24 subjects • Vehicle: 24 subjects	36 weeks	• Clinical assessment with high-resolution digital imaging of eye wrinkles, fine lines, brown spots, and skin-tone evenness (12-cm visual analog scale); three-dimensional profilometry of crow's feet area	• Statistically significant improvements ($p<0.05$) in retinol versus vehicle in eye wrinkles as of Week 12 and in fine lines as of Week 24	• Subjective and objective outcome measures but profilometry results not provided • Absolute data presented, including baseline values	• No primary endpoint designation • No power calculation • Unclear whether analysis is based on intention-to-treat or completers • No adjustment for multiple comparisons • No brown-spot or profilometry results
Kim et al.[13] (Jadad score: 4)	• Split-face design • Retinyl retinoate 0.06%/ Lotion P: 24 subjects	12 weeks	• Clinical assessment of periorbital wrinkles using a 0-7 scoring system, subject self-assessment, and visiometry of crow's feet area	• Statistically significant improvemaent ($p=0.031$) retinyl retinoate versus placebo in periorbital wrinkles as of Week 12	• Subjective and objective outcome measures provided but visiometry results not provided • Absolute data presented, including baseline values, with (?standard) deviations and p-values but no numbers of observations	• No primary endpoint designation • No power calculation • Unclear whether analysis is based on intention-to-treat or completers • No adjustment for multiple comparisons • No visiometry results

TABLE 3 (continued). Randomized, double-blind*, vehicle-controlled, whole- or split-face design trials of retinols in facial skin aging

STUDY (JADAD SCORE)	INTERVENTION WITH NUMBER OF SUBJECTS	TREATMENT DURATION	OUTCOME MEASURES	RESULTS	STRENGTHS	WEAKNESSES/LIMITATIONS
Lee et al.[14] (Jadad score: N/A)	• Split-face design • Retinyl N-formyl aspartamate/ vehicle: 24 subjects	24 weeks	• Clinical assessment and subject self-assessment of crow's feet area • Rating scale: 0, no improvement; 1, mild improvement; 2, moderate improvement; 3, remarkable improvement • Visiometry	• Statistically significant improvement retinyl N-formyl aspartamate versus vehicle in crow's feet area as of Week 24; visiometry showed statistically significant improvement (p=0.0014) versus vehicle for skin roughness as of Week 24	• Subjective and objective outcome measures Bonferroni adjustment for multiple comparisons • Absolute data presented for visiometry, including baseline values, standard deviations, and p-values	• No primary endpoint designation • No power calculation • No baseline information for clinical assessment • Rater not blinded as to treatment (single-blind) • Completer analysis
Randhawa et al.[25] (Jadad score: 3)	• Whole-face design • Retinol 0.1%: 35 subjects • Vehicle: 32 subjects	52 weeks	• Clinical assessment with photography of eight skin parameters, using a scoring system of 0 to 9, and subject self-assessment	• Statistically significant improvements ($p \leq 0.05$) retinol versus vehicle in all eight skin parameters as of Week 24	N/A	• No primary endpoint designation • No power calculation • Randomization unclear • No baseline information • Unclear whether analysis is based on intention-to-treat or completers • No adjustment for multiple comparisons • No results presented for vehicle • Retinol results presented without standard deviations or confidence intervals
Gold et al.[22] (Jadad score: 3)	• Whole-face design • Retinol 0.5%: 22 subjects • Vehicle: 8 subjects	8 weeks	• Clinical assessment of photodamage, wrinkles, ultraviolet spots, and photoaging and subject self-assessment of facial skin quality	• No statistically significant differences in retinol versus vehicle for all outcome measures	• Bonferroni adjustment for multiple comparisons	• No primary endpoint designation • No power calculation • No baseline information

*The trial by Lee et al.[14] is was a single-blind study, i.e., the subjects but not the evaluator(s) were blinded.

allowing for evaluator bias. An intention-to-treat analysis was not performed in any of the trials, thus potentially overestimating the treatment benefit by eliminating those subjects who dropped out after randomization and treatment initiation. Subjects generally drop out after the initiation of treatment because of negative experiences related to efficacy or side effects, unduly affecting the treatment outcome in a favorable way ("completer analysis"). In terms of statistical analysis, five trials performed multiple comparisons but only two corrected for them, increasing the likelihood of false-positive findings. Only three of the nine trials provided both baseline data and the results of treatment with *p*-values, while the others provided only one or the other or neither.

Four of the trials used a split-face design (side-to-side comparison), with no details provided as to how the risk of cross-contamination would be controlled. The use of a split-face design to compare topical treatments is efficient from a subject perspective but comes with the risk of subjects mixing up treatments, invalidating the results.[15]

Endpoint assessment in the trials was done by clinical evaluation, profilometry, or both. Five trials performed profilometry and only one (Lee et al.[14]) reported a statistically significant improvement versus vehicle in signs of (photo-)aging. Two of the trials did not provide the results of the profilometry that was described as part of the trials in the methods section. Using profilometry, Lee et al.[14] observed a statistically significant improvement versus vehicle in skin roughness only, which is interesting considering the fact that skin roughness improvement is not commonly documented in clinical trials of tretinoin.[5]

It is well established that, even under optimal conditions of subject preparation and lighting, real-time or photographic clinical evaluation can be difficult because even a seemingly minor change in facial expression or lighting can dramatically alter the appearance of facial skin lines and wrinkles.[16] Profilometry, on the other hand, offers an objective and quantitative method

TABLE 4. Key requirements for clinical trials to provide valid and meaningful results

REQUIREMENT	DEFINITION	IMPORTANCE
Demographic and baseline disease characteristics	Relevant demographic and disease characteristics after randomization but before the initiation of treatment should be presented for the subjects as a whole and separately for the treatment groups to be compared.	The treatment groups to be compared should be similar in terms of demographic and baseline disease characteristics. It confirms that randomization was successful in eliminating baseline differences between the groups, providing treatments with an equal start.
Randomization of subjects to treatment	Randomization is the deliberately haphazard method of allocating subjects to treatment.	The goal of randomization is to ensure that the treatment groups to be compared are as similar as possible, providing treatments with an equal start.
Blinding of subjects and investigators	Blinding concerns the concealment of treatment allocation and should be secured for subjects and investigators to render a double-blind trial design.	Without the blinding of subjects and investigators regarding treatment, observations will be biased—more so with subjective outcomes, such as clinical evaluation, than with objective outcomes, such as profilometry, but, in principle, to some degree with both.
Primary endpoint designation	Primary endpoint is the main trial outcome measure to evaluate the impact of the treatment.	The primary endpoint is the basis of the power calculation and sample size determination.
Power calculation	Power calculation relates to the probability of making a type II error—that is, not finding a difference between treatments while, in fact, there is one (i.e., false negative).	Without a power calculation, it is not possible to know whether a positive outcome is due to chance or a negative outcome is due to lack of power of the trial to detect a difference.
Correction for multiple comparisons	Correction for multiple comparisons controls the probability of a type I error—that is, finding a difference while, in fact, there is none (i.e., false positive).	Without correction for multiple comparisons, the likelihood of false positives—that is, finding a difference when, in fact, there is none—increases with the number of comparisons made.
Intention-to-treat analysis	In an intention-to-treat analysis, all randomized subjects are included in the analysis; in contrast, in a completer analysis, only the subjects who completed the trial are included.	An intention-to-treat analysis, as opposed to a completer analysis, also includes the subjects who dropped out after treatment initiation, which is generally caused by negative experiences related to efficacy or side effects. Not including the subjects who dropout potentially increases the study effect size—that is, the difference between the treatments compared.

to evaluate the effects of topical treatments on the signs of (photo-)aging. In profilometry, the surface topography of the skin is replicated using a silicone rubber impression, after which digital image processing is used to quantify the "microtopographic" features of the skin. The results are commonly reported as Rz, a measure of deeper wrinkles, and Ra, a measure of fine wrinkles and overall skin topography, respectively.[16] Probably the most reliable way to assess skin changes in clinical trials is a combination of blinded clinical evaluation and profilometry. Only five of the nine identified trials performed both, with two of them not reporting the profilometry results probably because no statistically significant differences versus vehicle were observed. The association between clinical evaluation and profilometry versus negative and positive outcomes is shown in Table 4.

The effect of vehicle or placebo in clinical trials is well-recognized and can be relatively high with topical treatments, similar to what is observed in analgesic or antidepressant trials.[17] It must be remembered that trials of topical treatments can only be vehicle-controlled; they cannot be placebo-controlled.[18] The vehicles used are almost certainly going to have some occlusive or emollient properties and, hence, are "moisturizers" known to have temporary as well as potentially longer-term physiological effects on the skin.[19] In the study by Kim et al.,[13] the vehicle response was virtually zero over a treatment period of 12 weeks and, in the study by Bellemère et al.,[20] the vehicle also showed virtually no change from baseline over the 36 weeks of the trial. In the study by Tucker-Samaras et al.,[21] of the 10 outcome measures assessed, there was no response to vehicle in six of them at Week 4 and five of them at Week 8, while all but one outcome measure (skin sagging) improved significantly relative to baseline in the treatment group. A placebo effect of zero with the use of a moisturizing vehicle over 8 to 36 weeks of treatment is not congruent with what would be expected and, hence, calls into question the reliability of the assessments or methods thereof.

The trials by Tucker-Samaras et al.[21] and Gold et al.[22] had the shortest treatment periods of eight weeks only. Although Gold et al.[22] used a retinol 0.5%–containing product and Tucker-Samaras et al.[21] used a retinol 0.1%–containing product, Gold et al. found no statistically significant differences from vehicle, while Tucker-Samaras et al. did. Aside from the obvious difference in retinol strengths, trial evidence suggests that topical tretinoin treatment improves the clinical appearance of (photo-)aged skin only after three to six months of treatment,[23] with many showing most of the clinically noticeable improvements to occur after six months.[3] Eight weeks of treatment as was adopted in the Tucker-Samaras et al. trial would be a surprisingly short period of time to allow for actual quantifiable improvement to be seen with a much less potent topical.

Finally and regarding a potential conflict of interest, eight of the nine trials were sponsored by the test-product manufacturer. It is unclear whether the trial published by Lee et al.[14] was industry-sponsored. Three (Tucker-Samaras et al.,[21] Kikuchi et al.,[24] and Randhawa et al.[25]) of the eight industry-sponsored trials were sponsored by the same global skincare company. Though the test product was not the same in the three trials (retinol 0.1% moisturizer, retinol 0.075% cream, and "stabilized" retinol 0.1%), the trials had the "strongest" positive results compared to the other trials and similar methodological flaws. Whether this is merely coincidence, a

true reflection of efficacy, or systematic bias is unclear.

CONCLUSION

Clinical trials often have results that are more impressive than what is observed in real life because of subjects' adherence to treatment, which is essential in trials and, therefore, emphasized and monitored.[26] In the real world, consumers pay for the products they use, which generally motivates them to use them, especially when they are expensive.[26] As demonstrated in numerous vehicle-controlled dermatological trials, the vehicle can facilitate impressive improvements on its own. It stands to reason that, if a patient is compliant with the treatment and has high expectations of it due to marketing and positive reviews, it is bound to "work" and meet the consumer's expectations, regardless of whether the active ingredient is indeed "active" or not. It is certainly possible that the improvements seen with cosmetic skincare products containing "active" ingredients are due to the vehicle alone.

An analysis of the nine clinical trials of retinol versus vehicle demonstrates this phenomenon. Four of the nine trials reported no statistically significant differences between the retinol-containing treatment and vehicle. The remaining five trials have major methodological flaws affecting design and analysis, which calls into question any positive results obtained. If cosmetic manufacturers are keen to show true efficacy for their retinol-containing products, they would be wise to adhere to the generally accepted guidelines for randomized, double-blind, vehicle-controlled trials and the publication thereof. This includes primary endpoint designation; power calculation; sample size determination; intention-to-treat analysis; well-described randomization and double-blinding processes; reporting of subject demographic and baseline disease characteristics; reporting of all primary, secondary, and exploratory outcomes as absolute values or absolute or percentage changes from baseline, along with standard deviations or confidence intervals; and performing appropriate statistical analyses with correction for multiple comparisons, if applicable. Until at least one—although preferably more than one for confirmatory purposes—high-quality clinical trials of retinol-containing products in the treatment of (photo-)aged skin are published, there is very little, if any, trustworthy evidence available to support the use of retinol-containing products to improve the appearance of (photo-)aged skin.

TABLE 5. Relationship between the method of assessment (e.g., clinical evaluation, profilometry) and negative versus positive outcomes

	NEGATIVE	POSITIVE	INCONCLUSIVE
Clinical evaluation	• Green et al.[28] • Gold et al.[22]	• Tucker-Samaras et al.[21] • Kikuchi et al.[24] • Bellemère et al.[20] • Kim et al.[13] • Randhawa et al.[25]	• Lee et al.[14] (investigator not blinded)
Profilometry	• Green et al.[28] • Creidi et al.[27] • Bellemère et al.[20]* • Kim et al.[13]*	• Lee et al.[14]**	N/A

The existing positive trials do provide weak evidence for retinol potentially having a mild ameliorating effect on fine facial skin wrinkle lines only. One might ask, is this enough evidence to justify choosing a retinol over a prescription tretinoin or viewing retinols as "equivalent" to tretinoin? However, equivalence can only be established through active-control trials with a noninferiority design and, to date, no such trials for retinols have been performed. In medicine in general with treatments impacting morbidity, quality of life, and mortality, basing treatment recommendations on poor-quality clinical trials would be considered negligent practice. It can be suggested that, in the case of retinols, the "positive" trials discussed in this review are not appropriate to inform clinical decision-making but rather serve as tools for advertising and marketing. No one's life is at risk if these trials do not provide valid data and, because they constitute over-the-counter products, no physician-based clinical decision-making is involved. Hence, skincare companies are not motivated to invest the time, money, or effort required to design, execute, and analyse high-level clinical trials of their cosmetic skincare products. Instead, they rely on the relative ignorance of the consumer, using carefully worded and often misleading statements based on poor-quality trials to sell their products. It is thus imperative for physicians as well as other health care professionals involved in evidence-based dermatology practice to regularly appraise such trials critically, allowing them to guide consumers in their choice of products and treatments.

REFERENCES

1. Khalil S, Bardawil T, Stephan C, et al. Retinoids: a journey from the molecular structures and mechanisms of action to clinical uses in dermatology and adverse effects. *J Dermatolog Treat.* 2017;28(8):684–696.
2. Connor MJ, Smit MH. The formation of all-trans-retinoic acid from all-trans-retinol in hairless mouse skin. *Biochem Pharmacol.* 1987;36(6):919–924.
3. Mukherjee S, Date A, Patravale V, et al. Retinoids in the treatment of skin aging: an overview of clinical efficacy and safety. *Clin Interv Aging.* 2006;1(4):327–348.
4. LLC VPNA. RENOVA ® (TRETINOIN CREAM) 0.02% Package Insert.; 2000.
5. Kang S, Bergfeld W, Gottlieb AB, et al. Long-term efficacy and safety of tretinoin emollient cream 0.05% in the treatment of photodamaged facial skin: a two-year, randomized, placebo-controlled trial. *Am J Clin Dermatol.* 2005;6(4):245–253.
6. Kurlandsky SB, Xiao JH, Duell EA, et al. Biological activity of all-trans retinol requires metabolic conversion to all-trans retinoic acid and is mediated through activation of nuclear retinoid receptors in human keratinocytes. *J Biol Chem.* 1994;269(52):32821–32827.
7. Kang S, Duell EA, Fisher GJ, et al. Application of retinol to human skin in vivo induces epidermal hyperplasia and cellular retinoid binding proteins characteristic of retinoic acid but without measurable retinoic acid levels or irritation. *J Invest Dermatol.* 1995;105(4):549–556.
8. Scientific Commmittee on Soncusmer

Safety. Final Version of the Opinion on Vitamin A (Retinol, Retinyl Acetate and Retinyl Palmitate) and Corrigendum. Available at: https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_199.pdf. Accessed May 17, 2020.
9. Diana Draelos Z. Therapeutic moisturizers. *Dermatol Clin.* 2000;18(4):597–607.
10. Cosmetic Toiletry and Perfumery Association (CTPA). Confidence in Cosmetic Claims. Available at: https://www.ctpa.org.uk/file.php?fileid=3330. Accessed May 17, 2020.
11. A-Passioni Retinol Cream by Drunk Elephant | Cult Beauty. Available at: https://www.cultbeauty.co.uk/drunk-elephant-a-passioni-retinol-cream.html. Accessed May 17, 2020.
12. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17(1):1–12.
13. Kim H, Kim N, Jung S, et al. Improvement in skin wrinkles from the use of photostable retinyl retinoate: a randomized controlled trial. *Br J Dermatol.* 2010;162(3):497–502.
14. Lee MS, Lee KH, Sin HS, er al. A newly synthesized photostable retinol derivative (retinyl N-formyl aspartamate) for photodamaged skin: profilometric evaluation of 24-week study. *J Am Acad Dermatol.* 2006;55(2):220–224.
15. Samuel M, Brooke R, Hollis S, et al. Interventions for photodamaged skin. *Cochrane Database Syst Rev.* 2015;2015(6):CD001782.
16. Grove GL, Grove MJ, Leyden JJ, et al. Skin replica analysis of photodamaged skin after therapy with tretinoin emollient cream. *J Am Acad Dermatol.* 1991;25(2):231–237.
17. Evers AWM. Using the placebo effect: how expectations and learned immune function can optimize dermatological treatments. *Exp Dermatol.* 2017;26(1):18–21.
18. Feldman SR. Placebo response? *Psychosom Med.* 2010;72(5):505.
19. Draelos ZD. The science behind skin care: moisturizers. *J Cosmet Dermatol.* 2018;17(2):138–144.
20. Bellemère G, Stamatas GN, Bruère V, et al. Antiaging action of retinol: from molecular to clinical. *Skin Pharmacol Physiol.* 2009;22(4):200–209.
21. Tucker-Samaras S, Zedayko T, Cole C, et al. A stabilized 0.1% retinol facial moisturizer improves the appearance of photodamaged skin in an eight-week, double-blind, vehicle-controlled study. *J Drugs Dermatol.* 2009;8(10):932–936.
22. Gold MH, Kircik LH, Bucay VW, et al. Treatment of facial photodamage using a novel retinol formulation. *J Drugs Dermatology.* 2013;12(5):533–540.
23. Bhawan J, Gonzalez Serva A, Nehal K, et al. Effects of tretinoin on photodamaged skin: a histologic study. *Arch Dermatol.* 1991;127(5):666–672.
24. Kikuchi K, Suetake T, Kumasaka N, et al Improvement of photoaged facial skin in middle-aged Japanese females by topical retinol (vitamin A alcohol): a vehicle-controlled, double-blind study. *J Dermatolog Treat.* 2009;20(5):276–281.
25. Randhawa M, Rossetti D, Leyden JJ, et al. One-year topical stabilized retinol treatment improves photodamaged skin in a double-blind, vehicle-controlled trial. *J Drugs Dermatology.* 2015;14(3):271–276.
26. Ali SM, Brodell RT, Balkrishnan R, et al. Poor adherence to treatments: a fundamental principle of dermatology. *Arch Dermatol.* 2007;143(7):912–915.
27. Creidi P, Vienne MP, Ochonisky S, et al. Profilometric evaluation of photodamage after topical retinaldehyde and retinoic acid treatment. *J Am Acad Dermatol.* 1998;39(6):960–965.
28. Green C, Orchard G, Cerio R, et al. A clinicopathological study of the effects of topical retinyl propionate cream in skin photoageing. *Clin Exp Dermatol.* 1998;23(4):162–167. JCAD